*Nonparametric Statistics*
Vol. 16(3–4), June–August 2004, pp. 591–611

# BOOTSTRAP TEST FOR CHANGE-POINTS IN NONPARAMETRIC REGRESSION

I. GIJBELS* and A-C. GODERNIAUX

*Institut de Statistique, Université catholique de Louvain, 20 Voie du Roman Pays, B-1348 Louvain-la-Neuve, Belgium*

*(Received 11 March 2003; Revised in final form 18 September 2003)*

The objective of this article is to test whether or not there is an abrupt change in the regression function itself or in its first derivative at certain (prespecified or not) locations. The test does not rely on asymptotics but approximates the sample distribution of the test statistic using a bootstrap procedure. The proposed testing method involves a data-driven choice of the smoothing parameters. The performance of the testing procedures is evaluated via a simulation study. Some comparison with an asymptotic test by Hamrouni (1999) and Grégoire and Hamrouni (2002b) and asymptotic tests by Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000) is provided. We also demonstrate the use of the testing procedures on some real data.

*Keywords*: Bandwidth; Bootstrap; Cross-validation; Discontinuity points; Derivative; Least-squares fitting; Local polynomial approximation

## 1 INTRODUCTION

In many applications, it may appear that a regression function is smooth except at an unknown finite number of points where jump discontinuities may occur. Examples are the Nile data (Cobb, 1978), the Mine accident data (Jarrett, 1979), the penny thickness data (see *e.g.* Scott, 1992) and the Prague temperature data (see Horváth and Kokoszka, 1997). Moreover, the regression function might be smooth (continuous) but its derivative might show a jump discontinuity. An example here is the Motorcycle data set (see *e.g.* Härdle, 1990) where changes in the 'direction' of the acceleration variable (*vs.* time after impact) are 'suspected' and to be tested for. This example involves testing for a jump discontinuity in the first derivative. We will illustrate the testing procedures developed in this article on some of these data sets.

If jump discontinuities are present in an otherwise smooth regression function or its derivative, their locations can be estimated followed by the estimation of the unsmooth curve (or derivative curve). The literature on estimation of locations of such change-points is by now quite large. Kernel-based estimation methods have been studied by Hall and Titterington (1992), Müller (1992), Wu and Chu (1993a, b, c), Chu (1994), Speckman (1994) and Eubank and Speckman (1994), among others. Speckman (1995) and Cline, *et al.* (1995) studied fitting curves with features using semiparametric change-point methods. Local polynomial methods

* Corresponding author. E-mail: gijbels@stat.ucl.ac.be

ISSN 1048-5252 print; ISSN 1029-0311 online 
DOI: 10.1080/10485250310001626088

were used by McDonald and Owen (1986), Loader (1996), Horváth and Kokoszka (1997, 2002), Qiu and Yandell (1998), Spokoiny (1998) and Hamrouni (1999) and Grégoire and Hamrouni (2002a), among others. For wavelet-based methods see, for example, Wang (1995), Raimondo (1998) and Antoniadis and Gijbels (2002). Müller and Song (1997) and Gijbels *et al.* (1999, 2003) proposed two-step estimation procedures. Gijbels and Goderniaux (2004a, b) discussed practical choices of the bandwidth parameters in the two-step procedure of Gijbels *et al.* (1999).

When estimating curves one often makes the assumption that the curve is smooth (*i.e.* at least continuous), leading to a smooth estimate. If such an assumption is doubted one would estimate first the change-points and then adapt the estimated regression curve to the estimated change-points. This unsmooth estimate usually appears quite different from the smooth estimate. See, for example, Müller and Stadtmüller (1999) for several plots illustrating this point. Hence, an important issue when estimating a regression function is to know whether it is reasonable to assume that the regression function (and/or its derivative(s)) is (are) smooth. This calls for testing the null hypothesis of a smooth regression function.

In this article, we discuss a bootstrap procedure for this testing problem, which does not rely on asymptotic laws. This is in contrast with testing procedures available in the literature which rely on asymptotic distributions of the estimators involved. References dealing with tests based on asymptotic laws include Wu and Chu (1993a, b, c), Hamrouni (1999), Müller and Stadtmüller (1999), Dubowik and Stadtmüller (2000) and Gregoire and Hamrouni (2002b). The bootstrap testing procedure presented in this article uses the two-step estimation method of Gijbels *et al.* (1999) as a starting basis. The reasons for this choice are: the method achieves the optimal rate of convergence, it has a good finite sample performance (as has been shown by extensive simulation studies) and data-driven choices of the parameters have been studied (*i.e.* the method is fully data-driven). As a consequence, the bootstrap testing procedures are also fully data-driven, not leaving the reader with a difficult and crucial choice of some smoothing parameters.

Following Hamrouni (1999) and Grégoire and Hamrouni (2002b) we discuss two tests: a local test and a global test. The local test focuses on testing whether the regression function (or its derivative) has a jump discontinuity at a certain fixed (prespecified) point $x_0$. Therefore, the local test relies on available information on the location of a possible jump discontinuity. With the global test we test the (more general) null hypothesis of a smooth curve *vs.* the alternative of a curve with at least one jump discontinuity (at an unknown point). We provide bootstrap testing procedures for both testing problems, and this for testing for discontinuities in the regression function itself or in its first derivative. Here, the procedures presented can be easily generalized for testing for jump discontinuities in the higher order derivatives of the regression function but for clarity of presentation we restrict ourselves to the first-order derivative. See Gijbels and Goderniaux (2004b) for ideas about basic adaptations to the case of higher order derivative curves.

The article is organized as follows. In Section 2, we focus on the bootstrap tests for the regression function itself, whereas in Section 3, we describe how to extend the procedures to obtain bootstrap tests for the testing problem involving the first derivative of the regression curve. Section 4 prepares for a comparison of the proposed bootstrap tests with some available asymptotic tests, such as those provided by Hamrouni (1999) and Grégoire and Hamrouni (2002b) and by Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000). A simulation study, presented in Section 5, illustrates the performances of the bootstrap-based tests and compares them with the performances of the asymptotic tests. In that section, we also demonstrate the use of the bootstrap tests on some real data.

## 2 BOOTSTRAP TESTING PROCEDURES

### 2.1 Statistical Model

We assume that a sample of $n$ data pairs $\mathcal{X} = \{(X_1, Y_1), \ldots, (X_n, Y_n)\}$ is observed, generated from the model

$$Y_i = g(X_i) + \varepsilon_i, \quad 1 \le i \le n. \tag{1}$$

The design points $X_i$ are either regularly spaced on $I = [0, 1]$ or are the order statistics of a random sample from a distribution having a density $f$ supported on $I$. The errors $\varepsilon_i$ are independent and identically distributed with zero mean and finite variance. The function $g$ is unknown and is assumed to be smooth (*i.e.* continuous) except at a finite (unknown) number of points.

### 2.2 The Local Test

#### *2.2.1 The Testing Problem*

Let $x_0$ be a fixed prespecified point in the interval $]0, 1[$. Then the interest is to test

$$\begin{aligned} &H_0\colon g \text{ is continuous in the point } x_0 \\ vs. \quad &H_1\colon g \text{ has a jump discontinuity at the point } x_0. \end{aligned}$$

More formally, consider the testing problem

$$\begin{cases} H_0\colon g(x_0^-) = g(x_0^+) \\ H_1\colon g(x_0^-) \neq g(x_0^+), \end{cases} \tag{2}$$

where $g(x_0^+) = \lim_{t \downarrow x_0} g(t)$ and $g(x_0^-) = \lim_{t \uparrow x_0} g(t)$ denote the right-hand limit, and the left-hand limit respectively, of the function $g$ at the point $x_0$.

Therefore, here we know in advance where the possible discontinuity could be located (namely at the point $x_0$) and we want to test whether there is indeed a jump occurring at this point or not. In other words, we want to test if the size of the jump at $x_0$ is significantly different from zero or not. Since the location $x_0$ is known, we estimate the unknown regression function $g$ in model (1) by

$$\hat{g}_{\mathrm{US}}(x) = \begin{cases} \hat{g}_1(x) & \text{if } x \in [0, x_0] \\ \hat{g}_2(x) & \text{if } x \in \,]x_0, 1], \end{cases} \tag{3}$$

where $\hat{g}_1$ and $\hat{g}_2$ are nonparametric estimates of $g$ on the interval $[0, x_0]$ and $]x_0, 1]$, respectively. The subscript US in $\hat{g}_{\mathrm{US}}(\cdot)$ refers to the fact that the resulting estimator is possibly unsmooth at the point $x_0$. We choose to work with kernel-type regression estimators, such as for example the Nadaraya–Watson estimator (see Nadaraya, 1964; Watson, 1964) or the local linear estimator (see Fan and Gijbels, 1996). Such estimators require the choice of a kernel function $K$ as well as a bandwidth parameter. The nonparametric estimator $\hat{g}_\ell$, involves a global smoothing parameter $h_\ell$, $\ell = 1, 2$. For simplicity we restrict to the case $h_1 = h_2 = h$.

One could think of several approaches for choosing the smoothing parameter $h$ in a data-driven way. A cross-validation procedure for choosing $h$ consists of minimizing the cross-validation quantity

$$\mathrm{CV}(h) = \sum_{i=1}^{i_0} \{\hat{g}_1^{-i}(X_i; h) - Y_i\}^2 + \sum_{i=i_0+1}^{n} \{\hat{g}_2^{-i}(X_i; h) - Y_i\}^2, \tag{4}$$

with $i_0 = \max\{i\colon X_i \le x_0\}$, and where $\hat{g}_1^{-i}(\cdot)$ and $\hat{g}_2^{-i}(\cdot)$ denote the estimators $\hat{g}_1$ and $\hat{g}_2$ obtained by discarding the $i$th data point, on the interval $[0, x_0]$ and $]x_0, 1]$, respectively. The cross-validated bandwidth selector is then defined as

$$\hat{h}_{\mathrm{CV}} = \arg\min_h \mathrm{CV}(h). \tag{5}$$

A cross-validation procedure for selecting two possibly different smoothing parameters $h_1$ and $h_2$ would be as follows:

$$\hat{h}_{1,\mathrm{CV}} = \arg\min_{h_1} \sum_{i=1}^{i_0} \{\hat{g}_1^{-i}(X_i; h_1) - Y_i\}^2$$

$$\hat{h}_{2,\mathrm{CV}} = \arg\min_{h_2} \sum_{i=i_0+1}^{n} \{\hat{g}_2^{-i}(X_i; h_2) - Y_i\}^2.$$

Define by $\gamma$, the jump size at the possible jump point $x_0$, *i.e.*

$$\gamma = g(x_0^+) - g(x_0^-).$$

Testing problem (2) is then equivalent to

$$\begin{cases} H_0\colon \gamma = 0 \\ H_1\colon \gamma \neq 0. \end{cases}$$

We consider the test statistic

$$T = \hat{\gamma} = \hat{g}_2(x_0^+) - \hat{g}_1(x_0), \tag{6}$$

with $\hat{g}_1$ and $\hat{g}_2$ as in Eq. (3). The null hypothesis is rejected if we obtain a value for $|T|$ that is too large. Denoting by $T_{\mathrm{obs}}$ the observed value of the test statistic, we would conclude that the size of the jump at $x_0$ is significantly different from zero if

$$T_{\mathrm{obs}} > c_1\left(\frac{\alpha}{2}\right) \quad \text{or} \quad T_{\mathrm{obs}} < c_2\left(\frac{\alpha}{2}\right),$$

where $c_1(\alpha/2)$ and $c_2(\alpha/2)$ are the two $\alpha/2$-critical levels of the distribution of $T$, *i.e.*

$$c_1\left(\frac{\alpha}{2}\right) = \inf\left\{x\colon P_{H_0}\{T \le x\} \ge 1 - \frac{\alpha}{2}\right\} \quad \text{and} \quad c_2\left(\frac{\alpha}{2}\right) = \sup\left\{x\colon P_{H_0}\{T \le x\} \le \frac{\alpha}{2}\right\}.$$

In order to determine these critical points, we must know the distribution of the statistic $T$ under $H_0$. But this distribution is unknown and we propose to use a bootstrap procedure to approximate these two critical points.

### *2.2.2 The Bootstrap Procedure*

There are two important aspects in the bootstrap procedure. Firstly, we need to resample from the residuals $\varepsilon_i = Y_i - g(X_i)$, and for this we wish to mimic the 'true' $\varepsilon$s as closely as possible. Best estimation of these $\varepsilon$s is obtained via 'best' estimation of $g$ itself. Since $\varepsilon_i = Y_i - g(X_i)$, with $Y_i$ known, 'best' estimation of $\varepsilon_i$ reduces to 'best' estimation of $g$ itself. More precisely, choosing $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n$ to minimize $\sum(\hat{\varepsilon}_i - \varepsilon_i)^2$ is equivalent to choosing $\hat{g}$ to minimize $\sum\{\hat{g}(X_i) - g(X_i)\}^2$. Using $\hat{g}_{\mathrm{US}}$, defined in Eq. (3) to calculate the residuals seems to be a reasonable option. In case $H_0$ is not true, this is the option to choose, and even if $H_0$ is true the estimate $\hat{g}_{\mathrm{US}}$ will still provide a consistent estimate of $g$. Second, to approximate the distribution of the test statistic $T$ we wish to mimic the data-generating mechanism under the null hypothesis of continuity of $g$. Hence, in the bootstrapping stage below (see Step 2) we use an estimator $\hat{g}_{\mathrm{S}}$ of $g$, based on the data $(X_1, Y_1), \dots, (X_n, Y_n)$, obtained under the assumption that $g$ is a smooth function having no discontinuities (*i.e.* $\gamma = 0$). We opt here for using $\hat{g}_{\mathrm{S}}(\cdot; \hat{h}_{\mathrm{S,CV}})$, the local linear regression estimator with a cross-validated bandwidth $\hat{h}_{\mathrm{S,CV}}$ obtained via $\hat{h}_{\mathrm{S,CV}} = \arg\min_h \sum_{i=1}^n \{\hat{g}_{\mathrm{S}}^{-i}(X_i; h) - Y_i\}^2$.

For fixed design, the bootstrap procedure for assessing the critical points then reads as follows.

*Step 1 Computation of residuals*

Calculate the estimated residuals

$$\tilde{\varepsilon}_i = Y_i - \hat{g}_{\mathrm{US}}(X_i; \hat{h}_{\mathrm{CV}}),$$

with $\hat{h}_{\mathrm{CV}}$ as in Eq. (5). Let $\bar{\varepsilon}$ denote the mean of all $\tilde{\varepsilon}_i$s, and put $\hat{\varepsilon}_i = \tilde{\varepsilon}_i - \bar{\varepsilon}$.

*Step 2 Monte Carlo simulation*

Obtain $\varepsilon_1^*, \dots, \varepsilon_n^*$ by resampling with replacement from $\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n$. Obtain

$$Y_i^* = \hat{g}_{\mathrm{S}}(X_i; \hat{h}_{\mathrm{S,CV}}) + \varepsilon_i^*, \quad i = 1, \dots, n,$$

and get

$$T^* = \hat{g}_2^*(x_0^+) - \hat{g}_1^*(x_0), \tag{7}$$

where $\hat{g}_1^*(\cdot)$ and $\hat{g}_2^*(\cdot)$ are nonparametric estimates of $g$ on the intervals $[0, x_0]$ and $]x_0, 1]$, respectively, using the bootstrap sample $\mathcal{X}^* = \{(X_1, Y_1^*), \dots, (X_n, Y_n^*)\}$. To choose the bandwidths involved we use the cross-validation procedure (4).

*Step 3 Determination of the bootstrap critical points*

Repeat Step 2 a large number of times, say $B$ times, and approximate the critical points $c_1(\alpha/2)$ and $c_2(\alpha/2)$ by the smallest value $c_1^*(\alpha/2)$ and the largest value $c_2^*(\alpha/2)$ such that

$$\frac{\#\{b = 1, \dots, B\colon T_b^* \le c_1^*(\alpha/2)\}}{B} \ge 1 - \frac{\alpha}{2}$$
$$\frac{\#\{b = 1, \dots, B\colon T_b^* \le c_2^*(\alpha/2)\}}{B} \le \frac{\alpha}{2}.$$

Bowman *et al.* (1998) considered, in the context of testing for monotonicity of a regression function, a bootstrapping procedure similar in nature to the above. For some discussion of this and similar bootstrap procedures, in the context of bump hunting for regression, see also Harezlak (1998) and Harezlak and Heckman (2001).

Finally, the bootstrap test consists of rejecting $H_0$ if

$$T_{\text{obs}} > c_1^*\left(\frac{\alpha}{2}\right) \quad \text{or} \quad T_{\text{obs}} < c_2^*\left(\frac{\alpha}{2}\right).$$

In case of random design, the above bootstrap algorithm needs to be adjusted accordingly. This adjustment consists of also resampling from the design points $X_1, \ldots, X_n$ using smoothed bootstrap, and obtaining as such a bootstrap sample $\mathcal{X} = \{(X_1^*, Y_1^*), \ldots, (X_n^*, Y_n^*)\}$.

In Section 5.2 we compare, via simulations, the performance of this bootstrap test with the performance of an asymptotic test developed by Gregoire and Hamrouni (2002b).

## 2.3 The Global Test

### 2.3.1 The Testing Problem

Here we want to test whether the unknown regression function has at least one jump discontinuity or not. We do not specify in advance where a jump discontinuity might occur. Therefore, this situation is more appropriate when we have no idea in advance about the location of a possible jump discontinuity. On the basis of the observations $\{(X_1, Y_1), \ldots, (X_n, Y_n)\}$, we wish to test the null hypothesis

$$\begin{cases} H_0\colon g \text{ is continuous on the interval } ]0, 1[ \\ H_1\colon g \text{ is discontinuous in at least one point on the interval } ]0, 1[ \end{cases}$$

or, more formally

$$\begin{cases} H_0\colon \forall x_0 \in ]0, 1[\colon g(x_0^-) = g(x_0^+) \\ H_1\colon \exists x_0 \in ]0, 1[ \text{ such that } g(x_0^-) \neq g(x_0^+). \end{cases} \tag{8}$$

Note that the alternative hypothesis states that there is a discontinuity point at a point $x_0$, but there might be more than one jump discontinuity.

### *2.3.2 Bootstrap Algorithm and Testing Procedure*

For this more general testing problem we use the test statistic

$$T = \hat{\gamma} = \hat{g}_2(\hat{x}_0^+) - \hat{g}_1(\hat{x}_0),$$

which is similar to the test statistic defined in Eq. (6), but with the unknown $x_0$ replaced by the estimator $\hat{x}_0$, assuming that there is (at least) one jump discontinuity. We use the data-driven bootstrap method developed by Gijbels and Goderniaux (2004a) for estimating $x_0$. This method is based on a two-step estimation procedure introduced and studied by Gijbels *et al.* (1999, 2003), and consists of obtaining a preliminary estimator of $x_0$ defined as the location where a derivative estimate of $g$ achieves its maximum in absolute value. A refinement of this estimator is then obtained via least-squares fitting of a piecewise constant (or polynomial) function in a neighbourhood of the preliminary estimator.

Therefore, in order to calculate the statistic $T$, we first need to get $\hat{x}_0$ (via the data-driven two-step procedure) and then calculate the statistic $T$. The calculation of the bootstrapped version of $T$ also consists of these two stages: determination of the location of the possible jump discontinuity, followed by the estimation of the jump size at that location. This results into a bootstrap algorithm, similarly as before but now with Eq. (7) replaced by

$$T^* = \hat{g}_2^*(\hat{x}_0^{*+}) - \hat{g}_1^*(\hat{x}_0^*),$$

where $\hat{x}_0^*$ is the estimated location of the possible jump (the most important one) using the bootstrap sample.

In Section 5, we evaluate the performance of this global test via simulations for functions having one or two discontinuities and for fixed or random design.

## 3 GENERALIZATION FOR TESTING FOR DISCONTINUITIES IN DERIVATIVES

In this section, we show how to generalize the local and global tests for testing for discontinuities in the first derivative of the regression function. The generalization for testing for discontinuities in the $k$th derivative function ($k > 1$) is similar in spirit and is not presented here.

We continue to consider model (1) but we now assume that the function $g$ is continuous and its first derivative is continuous except at some finite (unknown) number of points.

### 3.1 The Local Test

For the local test we are interested in the testing problem

$$\begin{cases} H_0\colon g^{(1)}(x_0^-) = g^{(1)}(x_0^+) \\ H_1\colon g^{(1)}(x_0^-) \neq g^{(1)}(x_0^+), \end{cases}$$

where $g^{(1)}(x_0^+) = \lim_{t\downarrow x_0} g^{(1)}(t)$ and $g^{(1)}(x_0^-) = \lim_{t\uparrow x_0} g^{(1)}(t)$ denote the left-hand side limit and the right-hand side limit, respectively, of the first derivative of $g$ at the point $x_0$. In other words, we want to test if the first derivative $g^{(1)}$ of the regression function is continuous or not at a prespecified point $x_0$. This is equivalent to testing whether the size of the jump at $x_0$ in the first derivative is significantly different from zero or not.

Since the location $x_0$ is known we estimate the first derivative of the regression function $g^{(1)}$ in the model $Y_i = g(X_i) + \varepsilon_i$, by

$$\hat{g}_{\mathrm{US}}^{(1)}(x) = \begin{cases} \hat{g}_1^{(1)}(x) & \text{if } x \in [0, x_0] \\ \hat{g}_2^{(1)}(x) & \text{if } x \in ]x_0, 1], \end{cases}$$

where $\hat{g}_1^{(1)}(\cdot)$ and $\hat{g}_2^{(1)}(\cdot)$ are local linear estimates of $g^{(1)}$, the first derivative of the regression function $g$ on the interval $[0, x_0]$ and $]x_0, 1]$, respectively.

For choosing the bandwidth parameter in $\hat{g}_{\mathrm{US}}^{(1)}$ we use the cross-validation technique adapted for the estimation of a derivative proposed by Müller *et al.* (1987):

$$\begin{aligned} \mathrm{CV}^{(1)}(h) &= \sum_{i=1}^{i_0-1} \left\{ \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i} - \hat{g}_1^{(1),-(i,i+1)}(X_i^{(1)}, h) \right\}^2 \\ &\quad + \sum_{i=i_0}^{n-1} \left\{ \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i} - \hat{g}_2^{(1),-(i,i+1)}(X_i^{(1)}, h) \right\}^2, \end{aligned}$$

with $i_0 = \max\{i\colon X_i \leq x_0\}$ and $X_i^{(1)} = (X_{i+1} + X_i)/2$, and where $\hat{g}_1^{(1),-(i,i+1)}$ and $\hat{g}_2^{(1),-(i,i+1)}$ denote the estimates $\hat{g}_1^{(1)}$ and $\hat{g}_2^{(1)}$ obtained by discarding the $i$th and $(i+1)$th data points, on the interval $[0, x_0]$ and $]x_0, 1]$, respectively. The cross-validated bandwidth selector is then defined as $\hat{h}_{\mathrm{CV}} = \arg\min_h \mathrm{CV}^{(1)}(h)$. Similarly as before, a cross-validation procedure for selecting two separate smoothing parameters $h_1$ and $h_2$ would be as follows:

$$\hat{h}_{1,\mathrm{CV}} = \arg\min_{h_1} \sum_{i=1}^{i_0-1} \left\{ \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i} - \hat{g}_1^{(1),-(i,i+1)}(X_i^{(1)}; h_1) \right\}^2$$

$$\hat{h}_{2,\mathrm{CV}} = \arg\min_{h_2} \sum_{i=i_0+1}^{n-1} \left\{ \frac{Y_{i+1} - Y_i}{X_{i+1} - X_i} - \hat{g}_2^{(1),-(i,i+1)}(X_i^{(1)}; h_2) \right\}^2 .$$

As a test statistic we take the estimated jump size (in the first derivative) at the point $x_0$:

$$T_1 = \hat{\gamma}_1 = \hat{g}_2^{(1)}(x_0^+) - \hat{g}_1^{(1)}(x_0), \tag{9}$$

and reject the null hypothesis if the observed value of $|T_1|$ is too large. Approximated critical points of the test are obtained via a bootstrap procedure similar as the one presented in Section 2, but now with Eq. (7) replaced by

$$T_1^* = \hat{g}_2^{(1)*}(x_0^+) - \hat{g}_1^{(1)*}(x_0). \tag{10}$$

### 3.2 The Global Test

For the global test we are interested in the testing problem

$$\begin{cases} H_0\colon \forall x_0 \in ]0, 1[\colon g^{(1)}(x_0^-) = g^{(1)}(x_0^+) \\ H_1\colon \exists x_0 \in ]0, 1[ \text{ such that } g^{(1)}(x_0^-) \neq g(x_0^+). \end{cases}$$

As a test statistic we use

$$T_1 = \hat{\gamma}_1 = \hat{g}_2^{(1)}(\hat{x}_0^+) - \hat{g}_1^{(1)}(\hat{x}_0),$$

as defined in Eq. (9), but with $x_0$ replaced by an estimator $\hat{x}_0$. For this estimator we rely on the data-driven bootstrap method developed by Gijbels and Goderniaux (2004b), which in fact generalizes the method discussed in Gijbels and Goderniaux (2004a) for the regression function itself. In order to find approximated values, $c_1^*(\alpha/2)$ and $c_2^*(\alpha/2)$, of the two critical points of the test, we apply a bootstrap algorithm, similar to the one used for the local test but with Eq. (10) replaced by

$$T_1^* = \hat{g}_2^{(1)*}(\hat{x}_0^{*+}) - \hat{g}_1^{(1)*}(\hat{x}_0^*),$$

where $\hat{x}_0^*$ is the estimated location of the (possible) jump point occurring in the first derivative of the regression function based on the bootstrap sample. Note also that Step 1 of the bootstrap procedure is changed accordingly, since it now requires estimation of the unknown $x_0$ (estimated by $\hat{x}_0$) before the actual estimation of $\hat{g}_{\mathrm{US}}$ is done as indicated in Eq. (3).

## 4 ASYMPTOTIC TESTS

In Section 5, we provide some comparison of the bootstrap procedure proposed in Section 2 for testing for a continuous regression function *vs.* a noncontinuous regression function, with some asymptotic tests available in the literature. We consider the asymptotic tests suggested by Hamrouni (1999) and Grégoire and Hamrouni (2002b) and the testing procedures discussed in Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000). The last two papers focused on a quantity $\kappa$, 'the amount of discontinuity', which is defined as the sum of squared jump sizes at all jump points of the regression function, and proposed tests for testing $H_0$: $\kappa = 0$ *vs.* $H_1$: $\kappa > 0$. We now briefly describe these asymptotic testing procedures.

## 4.1 Test of Grégoire and Hamrouni

Hamrouni (1999) and Grégoire and Hamrouni (2002a, b) used local linear estimation for estimating a jump point $x_0$, and search for the location point $t$ for which the absolute difference between a right-hand estimate and a left-hand estimate of $g$ is maximal. More precisely, they consider

$$\hat{x}_0 = \arg\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|,$$

where $\hat{g}_-(\cdot)$ and $\hat{g}_+(\cdot)$ are local linear estimates of left-hand and right-hand limits of $g$. The jump size $\gamma(\gamma > 0)$ at the point $x_0$ is then estimated by $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$. Hamrouni (1999) established the asymptotic distribution for the statistic $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$. For the global testing problem, Hamrouni (1999) proposes a test that consists of rejecting $H_0$: '$g$ is continuous on the interval $]0, 1[$' when the observed value of $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$ is too large. The critical points for this test are based on the asymptotic law of $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$.

When the location $x_0$ of the jump point is known Hamrouni (1999) and Grégoire and Hamrouni (2002a, b) estimated the jump size at $x_0$ by $\hat{g}_+(x_0) - \hat{g}_-(x_0)$. For the local testing problem, Grégoire and Hamrouni (2002b) proposed two testing procedures, one based on the simple statistic $\hat{g}_+(x_0) - \hat{g}_-(x_0)$, and a second test based on a test statistic that takes into account the behaviour of $\hat{\gamma}(t) = \hat{g}_+(t) - \hat{g}_-(t)$ in a neighbourhood around $x_0$. Both testing procedures rely on the asymptotic distribution theory for the considered random quantities.

In Section 5, we provide some comparison of the performance of our bootstrap test with the performance of the strictly local test of Grégoire and Hamrouni (2002b) relying on the asymptotic law of $\hat{g}_+(x_0) - \hat{g}_-(x_0)$, and with the global test developed in Hamrouni (1999).

## 4.2 Tests of Dubowik, Müller and Stadtmüller

Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000) assumed that the data are generated from a fixed design regression model where the errors are i.i.d. with zero mean, finite variance $\sigma^2$ and finite fourth moment. The regression function is smooth except for an unknown number of jump points. The article focuses on estimating simultaneously the quantities $\sigma^2$, the error variance, and $\kappa$, the sum of squared jump sizes, and to test the null hypothesis $H_0$: $\kappa = 0$ ($g$ is continuous) *vs.* $H_1'$: $\kappa > 0$ ($g$ is discontinuous). The testing procedure is based on sums of squared differences of the data, formed with various span sizes:

$$Z_k = \sum_{j=1}^{n-L} \frac{(y_{j+k} - y_j)^2}{n - L} \quad 1 \le k \le L$$

where $L = L(n) \ge 1$ is a sequence of integers depending on $n$. Müller and Stadtmüller (1999) showed that the statistics $Z_k$ can be interpreted as dependent variables within the following three-parameter asymptotic linear model which contains the parameters of interest $\sigma^2$, $\kappa$ and

$\delta$, a parameter measuring the 'interaction' between the continuous and discontinuous part of the regression function:

$$Z_k = 2\sigma^2 + \frac{k}{n-L}\kappa + \left(\frac{k}{n-L}\right)^2 \delta + \tilde{\eta}_k \quad 1 \le k \le L,$$

with $\tilde{\eta}_k$ the error term.

Estimators for $\sigma^2$ and $\kappa$ are derived using least-squares and the asymptotic distribution for these estimators is established. This then is used to construct an asymptotic level $\alpha$ test for the null hypothesis of no change in the regression function. For more details about this procedure see Müller and Stadtmüller (1999). They also discuss fitting of the $Z_k$ via a simpler two-parameter (linear) model

$$Z_k = 2\sigma^2 + \frac{k}{n-L}\kappa + \eta_k \quad 1 \le k \le L.$$

From the simulation study provided in their paper they recommend using the three parameter linear model instead of the above simpler two parameter linear model. Dubowik and Stadtmüller (2000) proposed an improvement of the testing procedure based on the simpler linear model by symmetrizing the quantities $Z_k$. A similar improvement could probably be proposed when using the three-parameter linear model, provided the asymptotic distribution for the resulting estimators will be established.

## 5 SIMULATION STUDIES AND REAL DATA EXAMPLES

We now evaluate the finite sample performance of the proposed bootstrap testing procedures for the local and global testing problems. In Sections 5.1 and 5.2, we focus on tests for the regression function itself, whereas in Section 5.3 we illustrate the performance for the testing problem regarding the first derivative of the regression function. In Section 5.2, we compare the performances of the bootstrap-based test with the asymptotic (local and global) tests of Hamrouni (1999) and Grégoire and Hamrouni (2002b) on one hand and the asymptotic tests of Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000) on the other hand. Throughout the simulation study we took normally distributed errors with error variance $\sigma^2$. In Section 5.4, we apply the bootstrap test on the Prague temperature data and on the Motorcycle data.

### 5.1 Test for the Regression Function Itself

We consider five different models:

$$\begin{aligned}
g_1(x) &= 4x^2 + I(x > 0.5) \\
g_2(x) &= \cos\{8\pi(0.5 - x)\} - 2\cos\{(8\pi(0.5 - x)\}I(x > 0.5) \\
g_3(x) &= 4x^2 \\
g_4(x) &= \cos\{8\pi(0.5 - x)\} \\
g_5(x) &= 4x^2 + 1.2I(x > 0.2) + 0.8I(x > 0.5),
\end{aligned}$$

where $I(A)$ denotes the indicator function of the event $A$, *i.e.* $I(A) = 1$ if $A$ is true and zero otherwise. Figure 1 depicts these functions (as solid curves) together with a typical simulated

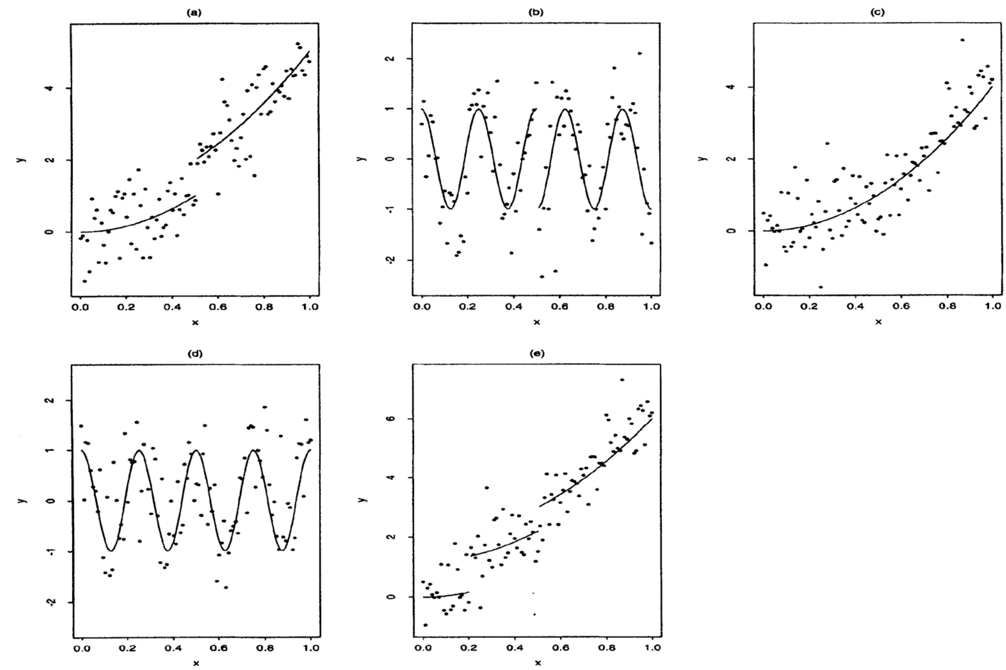


FIGURE 1 The true regression functions (solid curves) with a typical simulated data set of size $n = 100$ and variance $\sigma^2 = 0.5$. Regression functions: (a) $g_1$; (b) $g_2$; (c) $g_3$; (d) $g_4$ and (e) $g_5$.

dataset of size $n = 100$ for $\sigma^2 = 0.5$. The functions $g_1$ and $g_2$ have a jump at 0.5. The functions $g_3$ and $g_4$ are 'smooth' versions of $g_1$ and $g_2$, respectively. The function $g_5$, another 'discontinuous version' of the quadratic function $g_3$, has two discontinuity points occurring at 0.2 and 0.5.

### *5.1.1 Local Testing Problem*

To study the performance of the bootstrap test for the local testing problem

$$\begin{cases} H_0\text{: the function is continuous at the point 0.5} \\ H_1\text{: the function has a jump discontinuity at the point 0.5,} \end{cases}$$

we consider the models $g_1, g_2, g_3$ and $g_4$. We simulated samples of sizes $n = 50$, $n = 100$, $n = 250$ and $n = 800$ and took two different values of the error variance $\sigma^2 = 0.1$ and $\sigma^2 = 0.5$. We considered fixed equidistant design, $x_i = i/n$ for $i = 1, \ldots, n$ and we also dealt with random design, $X_i \sim U([0, 1])$ for the function $g_1$. For each setup we obtained 100 simulations and used $B = 2000$ bootstrap replications for the estimation of the two critical points of the test. The significance level is $\alpha = 0.05$. In Table I, we report on the proportion of times that the null hypothesis is rejected.

The functions $g_1$ and $g_2$ have a single jump discontinuity of size 1 and $-2$, respectively, at the point $x_0 = 0.5$. So the first two rows of Table I represent the proportion of times that $H_0$ was rejected when $H_0$ is false. Note that this proportion tends to one when $n$ increases, for $\sigma^2 = 0.1$ as well as for a bigger error variance $\sigma^2 = 0.5$. A similar result is seen for the function $g_1$ with random design (see the last row in Tab. I). These three rows illustrate the

TABLE I Results for the local test for the four different functions: the proportion of times $H_0$ was rejected.

| | | $n = 50$ | $n = 100$ | $n = 250$ | $n = 800$ |
|---|---|---|---|---|---|
| $g_1$ | $\sigma^2 = 0.1$ | 0.91 | 1 | 1 | 1 |
| Fixed design | $\sigma^2 = 0.5$ | 0.33 | 0.54 | 0.89 | 0.98 |
| $g_2$ | $\sigma^2 = 0.1$ | 0.95 | 1 | 1 | 1 |
| Fixed design | $\sigma^2 = 0.5$ | 0.53 | 0.76 | 0.95 | 1 |
| $g_3$ | $\sigma^2 = 0.1$ | 0.09 | 0.09 | 0.06 | 0.03 |
| Fixed design | $\sigma^2 = 0.5$ | 0.07 | 0.08 | 0.06 | 0.03 |
| $g_4$ | $\sigma^2 = 0.1$ | 0.06 | 0.07 | 0.05 | 0.04 |
| Fixed design | $\sigma^2 = 0.5$ | 0.06 | 0.06 | 0.09 | 0.04 |
| $g_1$ | $\sigma^2 = 0.1$ | 0.64 | 0.94 | 0.99 | 1 |
| Random design | $\sigma^2 = 0.5$ | 0.27 | 0.67 | 0.89 | 1 |

power properties of the bootstrap local test. The rows for the continuous functions $g_3$ and $g_4$ in Table I give an idea about the actual level of the bootstrap test.

### *5.1.2 Global Testing Problem*

We now present some simulation results on the global testing problem (8). We consider the functions $g_1$, $g_2$, $g_3$, $g_4$ and $g_5$. We simulated samples of size $n = 100$ and used $B = 500$ bootstrap replications for each estimation of the jump point and $B = 500$ bootstrap replications for estimating the two critical points of the test. Some results are presented in Table II that list the proportion of times (out of 100 simulations) that the null hypothesis was rejected.

Note from Table II that the results for the function $g_1$ are quite good. The results for the function $g_2$ are not so good as for the function $g_1$. This is simply because the function $g_2$ represents a more difficult case since it is essentially a cosinus function, showing many fluctuations. A typical difficulty inherent to this type of examples is the problem of identifying the jump point and distinguishing it from points with high absolute derivatives (points of steep increase or decrease). The results for the continuous functions $g_3$ and $g_4$ illustrate the actual level of the global test.

For the function $g_5$, the quadratic function with two discontinuities, the simulations for sample size $n = 100$ and $\sigma^2 = 0.1$ resulted into a proportion of rejection of $H_0$ (while $H_0$ is false) of 0.98. The algorithm more easily detected the discontinuity at the point 0.2. This is not surprising since the jump size at 0.2 is larger than that at 0.5. We also obtained good results for the function $g_1$ with $n = 100$ and $\sigma^2 = 0.1$ for random design $X_i \sim U([0, 1])$. In this setup, the percentage of rejection of $H_0$ was 78%.

TABLE II Results for the global test for four different functions for sample size $n = 100$: the proportion of times that $H_0$ was rejected.

| | $\sigma^2 = 0.1$ | $\sigma^2 = 0.5$ |
|---|---|---|
| $g_1$ | 0.98 | 0.66 |
| $g_2$ | 0.80 | 0.48 |
| $g_3$ | 0.04 | 0.03 |
| $g_4$ | 0.02 | 0.07 |

## 5.2 Comparison with Asymptotic Tests

### *5.2.1 Local Testing Problem*

We first compare the strictly local test proposed by Grégoire and Hamrouni (2002b), based on the statistic $\hat{g}_+(x_0) - \hat{g}_-(x_0)$, with the bootstrap-based test. Note that in this case the two testing procedures are based on the above simple statistic and the only difference between the two is that the test of Grégoire and Hamrouni (2002b) relies on asymptotic theory, whereas the bootstrap test relies on bootstrap approximations of the distribution of the test statistic.

The asymptotic law of the test statistic depends on the design density ($f(x) = 1$ in our setup) that we estimated via a kernel density estimate with a Gaussian kernel and a normal reference type of bandwidth parameter (see for example Silverman, 1986). We simulated samples from the regression model with functions $g_1$, $g_2$, $g_3$ and $g_4$ and error variance $\sigma^2 = 0.5$ for various sample sizes. The results displayed in Figure 2 present again the proportion of times (out of 100 simulations) that the null hypothesis was rejected. The bandwidth parameter for the estimation of the size of the jump was chosen via cross-validation based on minimizing the estimated mean integrated squared error for estimating $g$ on $[0, x_0]$ and $]x_0, 1]$. See also Eq. (4). This practical choice for the smoothing parameter was also suggested, but not further explored, in Grégoire and Hamrouni (2002b).

As we can see from Figure 2 the results for the two testing procedures, the asymptotic test and the bootstrap test, are very comparable. A better performance of the bootstrap-based test is noticable for small samples. See for example the function $g_4$ with $n = 50$ (the bottom left plot).

### *5.2.2 Global Testing Problem*

We now compare the global bootstrap test with the asymptotic tests for the global testing problem discussed briefly in Section 4. We will show results for the global test of Hamrouni (1999), for the test based on the three-parameter linear model for the $Z_k$s proposed by Müller and Stadtmüller (1999), and for the test of Dubowik and Stadtmüller (2000) based on the two-parameter linear model for the modified, symmetrized, $Z_k$s. We refer to the latter as the improved linear test.

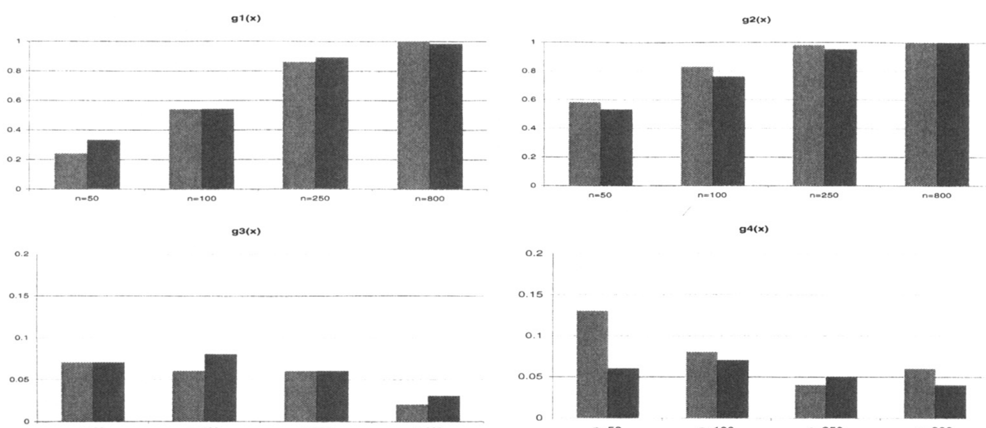


FIGURE 2 The proportion of times that $H_0$ was rejected for the bootstrap test (dark grey) and for the asymptotic test proposed by Grégoire and Hamrouni (light grey) for the regression functions $g_1$, $g_2$, $g_3$ and $g_4$ for different sample sizes and for error variance $\sigma^2 = 0.5$.

In this part of the simulation study, we considered six functions, the functions $g_1$, $g_2$, $g_3$, $g_4$ and $g_5$ defined in Section 5.1, and the function

$$g_6(x) = x + 2I(x > 0.5),$$

which was also considered in a simulation study in Müller and Stadtmüller (1999).

Recall that the functions $g_1$, $g_2$ and $g_6$ are discontinuous at the point 0.5, whereas the function $g_5$ has two discontinuity points occurring in 0.2 and 0.5. The functions $g_3$ and $g_4$ are continuous functions (*i.e.* $H_0$ is true).

The bootstrap test is fully data-driven. See Section 2.3. The global test proposed by Hamrouni (1999) requires calculation of the quantity $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$, using a bandwidth $h$, and the asymptotic theory establishes the asymptotic law for a normalized version of this statistic where the normalizing constants also depend on the bandwidth $h$. Hamrouni (1999) did not discuss practical choices of this smoothing parameter. In order to provide some comparison between the tests, we will, first, present the results for Hamrouni's test for a range of (fixed) values of $h$. The asymptotic tests proposed by Müller and Stadtmüller (1999) and Dubrowik and Stadtmüller (2000) involve the 'span parameter' $L$. Müller and Stadtmüller (1999) suggested a so-called 'plateau method' for selecting $L$ in practice. It is this method which we implemented in our simulations.

Figure 3(a)–(f) represents graphically the results of our simulations. On the horizontal axes the values of $h$ are presented, and the points in the plots indicate the proportion of times that $H_0$ was rejected for the Hamrouni (1999) test (for each of the fixed $h$-values). The long-dashed horizontal line indicates the result for the bootstrap test, whereas the short-dashed horizontal line represents the result for the asymptotic test of Müller and Stadtmüller (1999) with automatic choice of the parameter $L$. The dotted horizontal line shows the results for the improved linear test of Dubowik and Stadtmüller (2000). For cases for which the null hypothesis is true we also depict the significance level (0.05) as a solid horizontal line (see Fig. 3(a)–(b)). In Figure 3(a), the dotted and short-dashed line coincide and therefore only the dotted line is presented.

Figure 3 illustrates the impact of the smoothing parameter $h$ in the test of Hamrouni (1999). Further, we observe that for functions having one or two discontinuities (see Fig. 3(c)–(f)), the three different methods work quite well. The difference in performance of the three methods is most noticable for the cosinus functions $g_2$ and $g_4$. As already mentioned before, this is a rather difficult example since the cosinus function has many steep gradients that can blur the detection of the true discontinuity. To get some more inside in the impact of the span size parameter $L$ in the method of Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000) we calculated for the function $g_4$ (the cosinus function without a jump) the proportion of times $H_0$ was rejected, as a function of $L$. Figure 4 shows the simulated results, and indicates that for the continuous cosinus function, the choice of $L$ is very important, and the data-driven choice of $L$ does not work well for this example. The 'plateau method' for selecting $L$ focuses on the region where $L$ is reasonably stable (*i.e.* where the proportion stays constant), but apparently for this example the difficulty is that there are several such regions. A detailed investigation of our simulation results also revealed that for this example the estimation method of Hamrouni (1999), even when coming up with a bad estimation of the jump location, still estimates a reasonably big jump size at that (wrong) location. This can probably be explained by the steep increases and decreases of the cosinus function, and the identification problem for the jump discontinuity related to it. As a consequence the test too often rejects the null hypothesis. A similar observation was made from the simulations for the methods of Müller and Stadtmüller (1999) and Dubowik and Stadtmüller (2000). For the function $g_2$, the cosinus function with the jump, we noticed that the estimators of the sum of squared jump sizes often overestimates the true value 4. Again, a possible cause for this might be the rapid changes in the cosinus function.

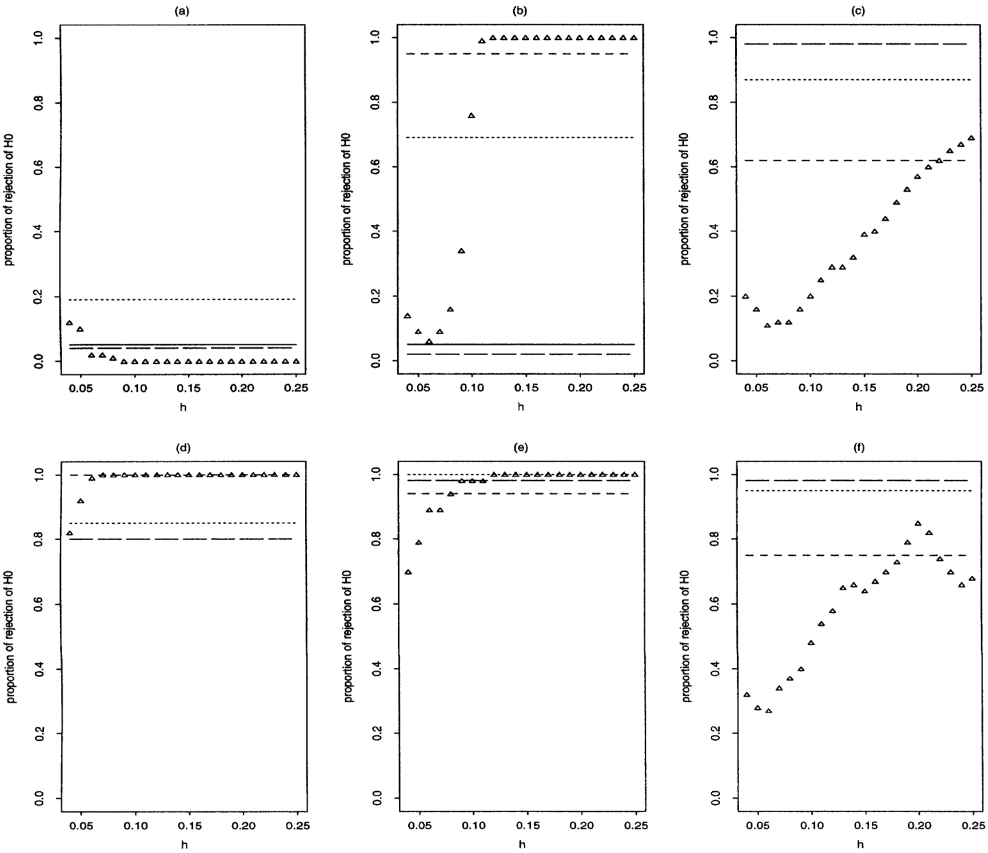


FIGURE 3 The proportion of times that $H_0$ was rejected for the bootstrap test (long-dashed horizontal line), for the Müller and Stadtmüller (1999) asymptotic test (short-dashed horizontal line), for the improved linear test of Dubowik and Stadtmüller (2000) (dotted horizontal line) and for the global test of Hamrouni (1999) (points indicated by the character '$\Delta$'). Results for the regressions functions: (a) $g_3$ (no jump); (b) $g_4$ (no jump); (c) $g_1$ (one jump); (d) $g_2$ (one jump); (e) $g_6$ (one jump); and (f) $g_5$ (two jumps).

This 'overestimation' of the jumps sizes can also be part of the explanation why these methods produce higher percentages of rejection for the function $g_4$.

One possible approach to choose the bandwidth parameter $h$ in the global test of Hamrouni (1999) would be to consider a large enough grid of $t$ values, and for each $t$ value choosing the cross-validation bandwidth which best estimates the curve $g$ on $[0, t]$ and on $]t, 1]$ (as in Eq. (4)). Suppose that the maximal quantity $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)|$ is then achieved at the value $t^*$. Then choose the bandwidth $h$ which corresponds to the cross-validation selected bandwidth for that $t^*$ value. This in fact means we opt for 'best' estimation of $\sup_t |\hat{g}_+(t) - \hat{g}_-(t)| \approx |\hat{g}_+(t^*) - \hat{g}_-(t^*)|$. We use this bandwidth in the test statistic as well as in the normalizing constant. The results for the Hamrouni test with this data-driven choice of the smoothing parameters are presented in Table III along with the previously obtained results on the other two tests. This table complements Figure 3.

## 5.3 Tests for the First Derivative

We now investigate the performance of the bootstrap test for testing whether the first derivative of the regression function is continuous or not. We consider three functions:

$$g_7(x) = 2x + 1 - 4(x - 0.5)_+,$$

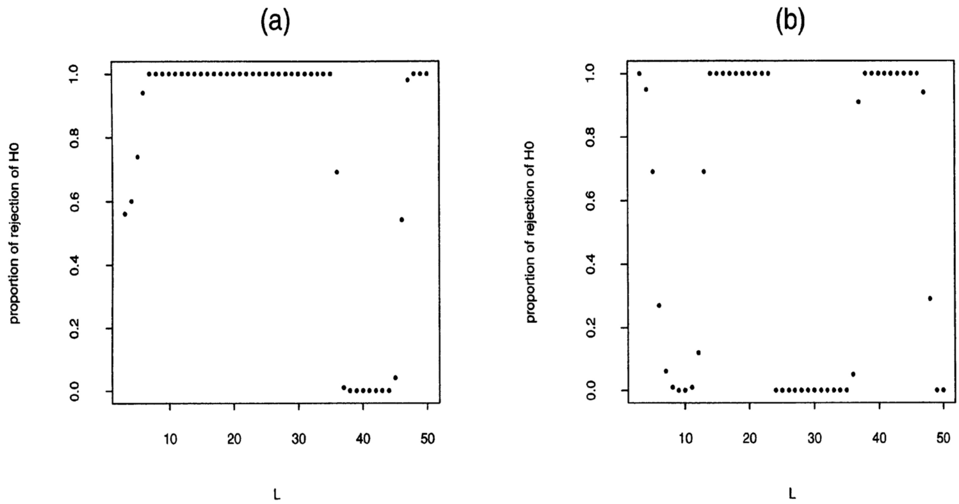


FIGURE 4 The proportion of times that $H_0$ was rejected for the asymptotic test of Müller and Stadtmüller (1999) (left panel) and the improved linear test of Dubowik and Stadtmüller (2000) (right panel), as a function of the span parameter $L$ for the continuous cosinus function $g_4$.

where $u_+ = \max(u, 0)$,

$$g_8(x) = \begin{cases} 10x^2 & \text{if } x \in [0, 0.5] \\ \dfrac{-20}{7}x^3 + \dfrac{20}{7} & \text{if } x \in ]0.5, 1] \end{cases}$$

and

$$g_9(x) = 10x^2.$$

The function $g_9$ and its derivative $g_9^{(1)}$ is smooth on $]0, 1[$, whereas the functions $g_7$ and $g_8$ have a first derivative that shows a discontinuity at the point 0.5. These three functions together with a simulated data set of size $n = 100$ and $\sigma^2 = 0.05$ are displayed in Figure 5 (upper panels). In the lower panels, we have represented the true first derivative with an adapted estimator in case $\sigma^2 = 0.01$ and $\sigma^2 = 0.05$. These pictures illustrate that estimation of derivatives curves is more difficult than estimation of the regression function itself. That is the reason why we consider smaller values of the error variance in this part of the simulation study.

TABLE III Results for the global test for the four different methods: the proportion of times that $H_0$ is rejected for the six different functions defined above. We use the sample size $n = 100$ with $\sigma^2 = 0.1$.

| | | *Hamrouni* | *Müller and Stadtmüller* | *Dubowik and Stadtmüller improved linear test* | *Bootstrap* |
|---|---|---|---|---|---|
| No jump | $g_3$ | 0.01 | 0.19 | 0.19 | 0.04 |
| | $g_4$ | 0.35 | 0.95 | 0.69 | 0.02 |
| One jump | $g_1$ | 0.57 | 0.62 | 0.87 | 0.98 |
| | $g_2$ | 0.99 | 1 | 0.85 | 0.80 |
| | $g_6$ | 1 | 0.94 | 1 | 1 |
| Two jumps | $g_5$ | 0.85 | 0.75 | 0.95 | 0.98 |

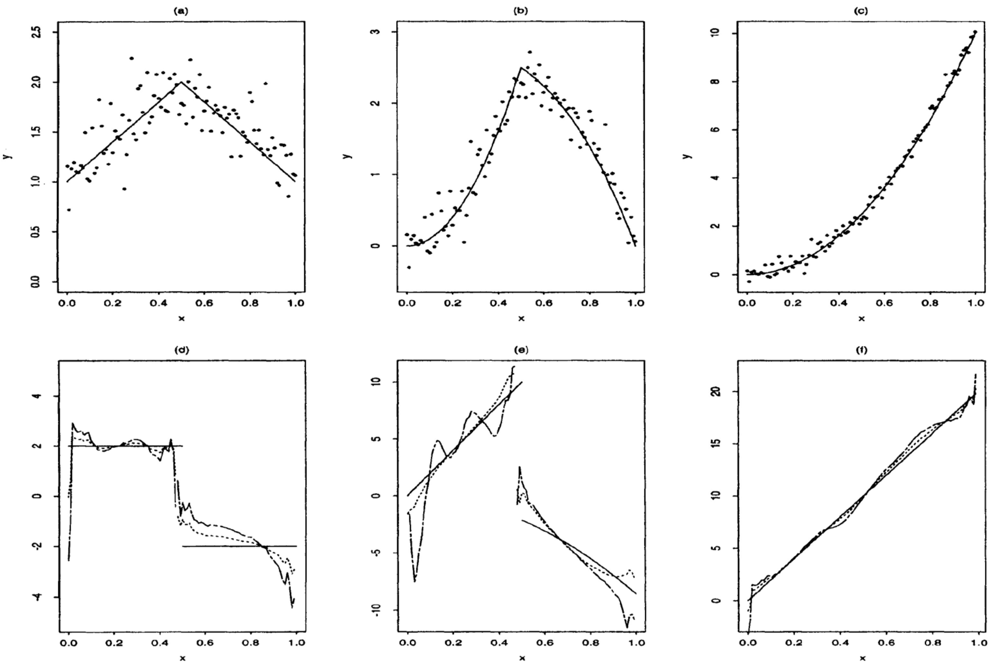


FIGURE 5 Upper panels: the true regression functions (solid curves) with a typical simulated data set of size $n = 100$ and with error variance $\sigma^2 = 0.05$. Regression functions: (a) $g_7$; (b) $g_8$ and (c) $g_9$. Lower panels: the true first derivative (solid curve) of the regression function with an estimator for the case $\sigma^2 = 0.01$ (dotted curve) and for the case $\sigma^2 = 0.05$ (dashed curve) for the regression function: (d) $g_7$; (e) $g_8$ and (f) $g_9$.

### *5.3.1 Local Testing Problem*

We simulated data of sizes $n = 50$, $n = 100$, $n = 250$ and $n = 800$ for two different values of the error variance, $\sigma^2 = 0.01$ and $\sigma^2 = 0.05$. For each setup we simulated 100 samples, and used $B = 2000$ bootstrap replications for approximating the two critical points of the test. We take $\alpha = 0.05$. Table IV lists the proportions of times (out of 100) that the null hypothesis ($H_0$: the first derivative is continuous at the point 0.5) was rejected. Since the functions $g_7$ and $g_8$ show a change in the derivative at 0.5, the rows for these functions indicate the percentage of rejecting $H_0$ when it is indeed false, and hence give an idea about the power of the test. Note that this percentage increases with $n$, indicating that the test is asymptotically efficient (power tending to 1). For the function $g_9$ the values presented in Table IV indicate the actual level of the test (which should be compared with the significance level $\alpha = 0.05$).

TABLE IV Results for the local test: the proportion of times that $H_0$: $g^{(1)}$ is continuous in the point 0.5 was rejected, for sample sizes $n = 50$, $n = 100$, $n = 250$ and $n = 800$ for error variance $\sigma^2 = 0.01$ or 0.05.

| | | $n = 50$ | $n = 100$ | $n = 250$ | $n = 800$ |
|---|---|---|---|---|---|
| $g_7$ | $\sigma^2 = 0.01$ | 0.68 | 0.75 | 0.82 | 0.91 |
| | $\sigma^2 = 0.05$ | 0.39 | 0.53 | 0.63 | 0.75 |
| $g_8$ | $\sigma^2 = 0.01$ | 0.85 | 0.90 | 1 | 1 |
| | $\sigma^2 = 0.05$ | 0.48 | 0.60 | 0.69 | 0.97 |
| $g_9$ | $\sigma^2 = 0.01$ | 0.08 | 0.03 | 0.03 | 0.04 |
| | $\sigma^2 = 0.05$ | 0.07 | 0.04 | 0.04 | 0.02 |

#### *5.3.2 Global Testing Problem*

We now investigate the performance of the global test

$$\begin{cases} H_0\colon g^{(1)} \text{is continuous on the interval } ]0,1[ \\ H_1\colon g^{(1)} \text{is not continuous on the interval } ]0,1[ \end{cases}$$

for the functions $g_7$, $g_8$ and $g_9$ for sample size $n = 100$ and error variance $\sigma^2 = 0.01$. The number of bootstrap replications involved in the estimation of the location of the possible jump point is 500. We also used 500 bootstrap replications for approximating the two critical points of the test. The results are reported in Table V.

### 5.4 Real Data Examples

#### *5.4.1 The Prague Temperature Data*

In order to illustrate the performance of the global test involving the regression function itself, we applied the bootstrap test procedure to a real data set. The application concerns 215 average annual temperatures measured in Prague from 1775 to 1989. This data set was analyzed by Horváth and Kokoszka (1997) in order to detect climatic changes that took place over a span of several years or a decade. We used 500 bootstrap replications for the estimation of the jump point and 500 for estimating the two critical points of the test. The estimated two critical points of the test are $c_1^*(0.025) = 2.39$ and $c_2^*(0.025) = -2.7$, and the observed value of the test statistic is $T_{\text{obs}} = 3.275$. Since $T_{\text{obs}} > c_1^*(0.025)$, we reject the null hypothesis that the regression function is continuous on $]0, 1[$. Note that the estimated $p$-value is equal to 0.0008. Gijbels and Goderniaux (2004a) analyzed this data set in order to estimate the number and the locations of the jump discontinuities and found three jump points occurring in 1786.5, 1836.5 and 1942.5. The data set with a local linear estimator adapted to these estimated change-points is presented in Figure 6. It should be mentioned that the estimated change-point 1786.5 is very close to the boundary of the observations, and hence estimation of $g$ between the first observation 1775 and this estimated change-point is not very accurate. The global bootstrap test detects the jump with the largest jump size *i.e.* the jump located at 1786.5.

#### *5.4.2 The Motorcycle Data*

We applied the global test for testing for discontinuities in the first derivative of the regression function to the Motorcycle data. This data set consists of 132 observations made on cadavers in

TABLE V Results for the global test: the proportion of times $H_0$ is rejected for the functions $g_7$, $g_8$ and $g_9$, for sample size $n = 100$ and error variance $\sigma^2 = 0.01$.

| | $\sigma^2 = 0.01$ |
|---|---|
| $g_7$ | 0.63 |
| $g_8$ | 0.84 |
| $g_9$ | 0.06 |

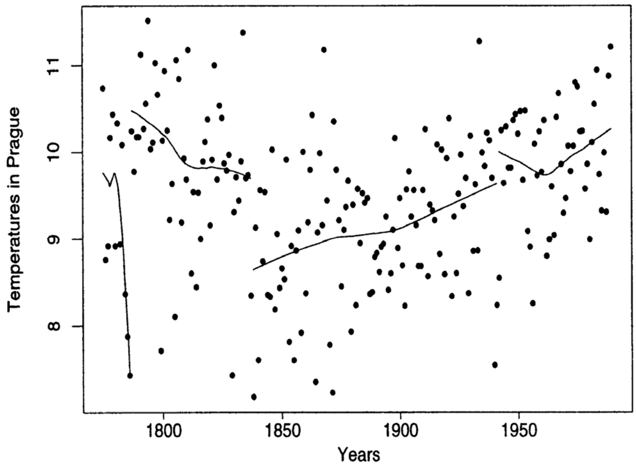


FIGURE 6 The average annual temperatures in Prague from 1775 to 1989 with a local linear estimator adapted to estimated discontinuities (solid curve).

simulated motorcycle collisions. The explanatory variable $X$ is the time (in ms) after impact, whereas the dependent variable $Y$ is the head acceleration (in $g$) of the test object. This data set has been analyzed by Speckman (1995) and Gijbels and Goderniaux (2004b).

The two critical points are estimated by the bootstrap procedure. We used 500 bootstrap replications for the estimation of the (most important) jump point and 500 for estimating the two critical points of the test. We obtained the following results: $c_1^*(0.025) = 10.88$ and $c_2^*(0.025) = -10.62$ which led to a rejection of the null hypothesis of a smooth first derivative since $T_{1,\text{obs}} = 91.89$. We then concluded that the first derivative of the regression function is not continuous. Figure 7 depicts the raw data and a local linear fit adapted to the three estimated change-points, 14.2, 21.1 and 32.4, given by Gijbels and Goderniaux (2004b).

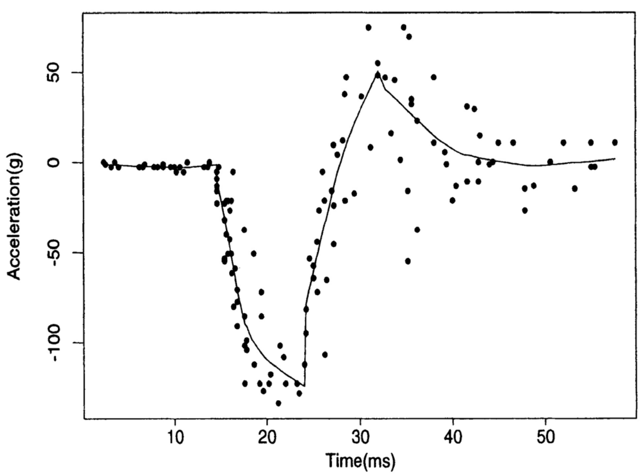


FIGURE 7 The Motorcycle data (points) with a local linear fit adapted to change-points (solid curve).

## Acknowledgements

The authors thank Professor Ulrich Stadtmüller for making them aware of the work by Dubowik and Stadtmüller. The authors also owe thanks to an anonymous referee for careful reading of the first version of the manuscript. Financial support from the contract 'Projet d'Actions de Recherche Concertées' nr 98/03-217 from the Belgian government, and from the IAP research network nr P5/24 of the Belgian State (Federal Office for Scientific, Technical and Cultural Affairs) is gratefully acknowledged.


## References


Antoniadis, A. and Gijbels, I. (2002). Detecting abrupt changes by wavelet methods. *Journal of Nonparametric Statistics*, **14**, 7–29.

Bowman, A. W., Jones, M. C. and Gijbels, I. (1998). Testing monotonicity of regression. *Journal of Computational and Graphical Statistics*, **7**, 489–500.

Chu, C. K. (1994). Estimation of change-points in a nonparametric regression function through kernel density estimation. *Communications in Statistics, Part A-Theory and Methods*, **23**, 3037–3062.

Cline, D. B. H., Eubank, R. and Speckman, P. (1995). Nonparametric estimation of regression curves with discontinuous derivatives. *Journal of Statistical Research*, **29**, 17–30.

Cobb, G. W. (1978). The problem of the Nile: conditional solution to a change-point problem. *Biometrika*, **62**, 243–251.

Dubowik, Ch. and Stadtmüller, U. (2000). Detecting jumps in nonparametric regression. In: *Asymptotics in Statistics and Probability*, Ed. Puri, M. L. pp 171–184.

Eubank, R. L. and Speckman, P.L. (1994). Nonparametric estimation of functions with jump discontinuities. In: *Change-point Problems*, Eds. Carlstein, E., Müller, H.-G., and Siegmund, D. IMS-Lecture Notes, Vol. 23, 130–143.

Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and its Applications*. Chapman and Hall, New York.

Gijbels, I. and Goderniaux, A-C. (2004a). Bandwidth selection for change point estimation in nonparametric regression. *Technometrics*, **46**, 76–86.

Gijbels, I. and Goderniaux, A-C. (2004b). Data-driven discontinuity detection in derivatives of a regression function. *Communications in Statistics, Part A – Theory and Methods*, **33**, 851–871.

Gijbels, I., Hall, P. and Kneip, A. (1999). On the estimation of jump points in smooth curves. *The Annals of the Institute of Statistical Mathematics*, **51**, 231–251.

Gijbels, I., Hall, P. and Kneip, A. (2003). Interval and band estimation for curves with jumps. *Journal of Applied Probability*. (Special Issue in honour of Chris Heyde), (to appear.).

Grégoire, G. and Hamrouni, Z. (2002a). Change-point estimation by local linear smoothing. *Journal of Multivariate Analysis*, **83**, 56–83.

Grégoire, G. and Hamrouni, Z. (2002b). Two nonparametric tests for change-point problem. *Journal of Nonparametric Statistics*, **14**, 87–112.

Hall, P. and Titterington, D. M. (1992). Edge preserving and peak-preserving smoothing. *Technometrics*, **34**, 429–440.

Hamrouni, Z. (1999). Inférence statistique par lissage linéaire local pour une fonction de régression présentant des discontinuités. *Doctoral thesis*, Université de Joseph Fourier, Grenoble, France.

Härdle, W. (1990). *Applied Nonparametric Regression*. Econometric Society Monograph Series Vol. 19, Cambridge University Press.

Harezlak, J. (1998). Bump hunting revisited. *MSc thesis*, Department of Statistics, University of British Columbia, Vancouver, Canada.

Harezlak, J. and Heckman, N. (2001). Crisp: a tool in bump hunting. *Journal of Computational and Graphical Statistics*, **10**, 713–729.

Horváth, L. and Kokoszka, P. (1997). Change-point detection with nonparametric regression. *Technical report*, University of Liverpool.

Horváth, L. and Kokoszka, P. (2002). Change-point detection with non-parametric regression. *Statistics*, **36**, 9–31.

Jarrett, R. G. (1979). A note on the intervals between coal-mining disasters. *Biometrika*, **66**, 191–193.

Loader, C. R. (1996). Change point estimation using nonparametric regression. *The Annals of Statistics*, **24**, 1667–1678.

McDonald, J. A. and Owen, A. B. (1986). Smoothing with split linear fits. *Technometrics*, **28**, 195–208.

Müller, H.-G. (1992). Change-points in nonparametric regression analysis. *The Annals of Statistics*, **20**, 737–761.

Müller, H.-G. and Song, K.-S. (1997). Two-stage change-point estimators in smooth regression models. *Statistics & Probability Letters*, **34**, 323–335.

Müller, H.-G. and Stadtmüller, U. (1999). Discontinuous versus smooth regression. *The Annals of Statistics*, 27, 299–337.

Müller, H.-G., Stadtmüller, U. and Schmitt, T. (1987). Bandwidth choice and confidence intervals for derivatives of noisy data. *Biometrika*, **74**, 4, 743–749.

Nadaraya, E. A. (1964). On estimating regression. *Theory of Probability and Applications*, **9**, 141–142.

Qiu, P. and Yandell, B. (1998). A local polynomial jump-detection algorithm in nonparametric regression. *Technometrics*, **40**, 141–152.

Raimondo, M. (1998). Minimax estimation of sharp change points. *The Annals of Statistics*, **26**, 1379–1397.

Scott, D. W. (1992). *Multivariate Density Estimation. Theory, Practice and Visualization*. Wiley, New York.
Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. Chapman and Hall, London.
Speckman, P. L. (1994). Detection of change-points in nonparametric regression. (unpublished manuscript.)
Speckman, P. L. (1995). Fitting curves with features: semiparametric change-point methods. *Computing science and statistics*, 26, 257–264.
Spokoiny, V. G. (1998). Estimation of a function with discontinuities via local polynomial fit with an adaptive window choice. *The Annals of Statistics*, **26**, 1356–1378.
Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika*, **82**, 385–397.
Watson, G. S. (1964). Smooth regression analysis. *Sankhyā* Series A, **26**, 359–372.
Wu, J. S. and Chu, C. K. (1993a). Kernel type estimation of jump points and values of regression function. *The Annals of Statistics*, **21**, 1545–1566.
Wu, J. S. and Chu, C. K. (1993b). Modification for boundary effects and jump points in nonparametric regression. *Nonparametric Statistics*, **2**, 341–354.
Wu, J. S. and Chu, C. K. (1993c). Nonparametric function estimation and bandwidth selection for discontinuous regression functions. *Statistica Sinica*, **3**, 557–576.